中国科学: 信息科学 2017 年 第 47 卷 第 4 期: 492–506
SCIENTIA SINICA Informationis
论文

基于博弈论的重大公共活动安保策略设计算法

殷越[1,2*], 安波[1,3], 史忠植[1]

1. 中国科学院计算技术研究所智能信息处理重点实验室, 北京 100190, 中国
2. 中国科学院大学, 北京 100049, 中国
3. School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, Singapore
* 通信作者. E-mail: melody1235813@163.com

收稿日期: 2016–04–08; 接受日期: 2016–06–27; 网络出版日期: 2017–02–23
国家自然科学基金 (批准号: 61202212) 资助项目

摘要 重大公共活动, 比如大型赛事, 由于其参与人数众多, 影响力广泛, 一直是恐怖分子的重要攻击目标. 因此, 重大公共活动的安保问题也是各国政府必须面对的一项难题. 由于公共活动通常场地复杂, 参与者多样, 而安全部门可支配的安保资源有限, 如何最大限度地利用有限的资源保障活动安全进行成为了一项极具挑战的任务. 本文以博弈模型来描述重大公共活动的安保问题, 该模型既考虑了公共活动本身人流量与时间相关的特点, 也考虑了安全部门与潜在的恐怖分子的复杂的策略空间. 基于此模型, 本文研究了安保资源转移时间可忽略与转移时间不可忽略两种情况, 并分别提出算法 SCOUT-A (Scheduling seCurity resOurces in pUblic evenTs with no relocating delAy) 和 SCOUT-C (Scheduling seCurity resOurces in pUblic evenTs against Continuous strategy space) 来求解安保部门的最优策略. 实验证明, 本文提出的算法比已有的算法为安保部门带来更好的收益.

关键词 博弈论, 安全, 算法设计, 策略设计, 多智能体

1 引言

近年来, 恐怖活动日益猖獗, 袭击重大公共活动的事件时有发生. 例如 2013 年的 Boston 马拉松爆炸案, 就造成了重大的损失和恶劣的影响. 保障重大赛事的安全是各国的安全部门都亟需解决的问题. 然而, 利用有限的安全资源设计高效的重大赛事安保策略却面临着技术上的挑战, 主要原因有如下 3 点. 第一, 重大赛事中可能被攻击的目标的重要程度是随时间变化的. 例如, 在马拉松比赛中, 赛道的某一段可视为一个可能被攻击的目标. 而参赛选手和观众随着比赛进行而沿着赛道转移, 不同赛段的重要程度也随之变化. 这就要求安保资源随时间动态分配. 第二, 潜在的恐怖分子可能在任何时间发动攻击, 而安全部门也可以在任何时间转移资源, 这也就意味着双方的策略空间都无限大. 第三, 由

引用格式: 殷越, 安波, 史忠植. 基于博弈论的重大公共活动安保策略设计算法. 中国科学: 信息科学, 2017, 47: 492–506, doi: 10.1360/N112016-00088
Yin Y, An B, Shi Z Z. Designing game-theoretic security strategies for large public events (in Chinese). Sci Sin Inform, 2017, 47: 492–506, doi: 10.1360/N112016-00088

于重大赛事的相关信息通常会提前公布, 恐怖分子在制定攻击计划时也会策略性地考虑这些信息. 因此, 安保部门在设计安保策略时也需要考虑恐怖分子的策略. 本文致力于解决这 3 个问题, 以博弈模型描述安全部门和恐怖分子的策略行为, 并设计算法求解安全部门的最优策略.

博弈论已被成功应用于很多安全领域, 例如机场安保 [1]、港口巡逻 [2,3], 以及其他多种形式的安保策略设计 [4~7]. 然而, 已有的模型与算法并不能被直接用于解决重大赛事的安保问题. 首先, 大多数的研究都假设可能被攻击的目标的重要程度是不随时间变化的 [8~11], 这与重大赛事场地复杂的情况不符. 虽然有部分研究人员考虑了攻击目标的重要程度可能随时间变化的情况 [12], 他们的工作与本文有两点本质的区别. 其一, 他们假设安全部门和恐怖分子都只能在预先设定的某些时间点行动 [13]. 第二, 他们关注的是常规的、通常需要每天都执行的安保策略, 因而恐怖分子在行动时会参考安保部门已经执行过的策略. 而由于重大公共活动频率较低, 恐怖分子并没有关于安保策略的前期资料可参考, 其行为取决于其对赛事情况的了解以及基于此对于安保策略的预测. 因此, 在重大赛事的安保领域, 恐怖分子与安保部门直接的交手更应该被看做无历史信息的一次性行为 [14].

本文首先设计了一种博弈模型来描述重大公共活动的安保问题. 在此模型中, 安全部门 (以下简称安保方) 与恐怖分子 (以下简称攻击方) 均具有连续的策略空间. 在此博弈模型的均衡策略下, 即安保方的最优策略下, 安保方可能遭受的最大的损失最小. 接着, 本文关注了资源的转移时间远小于活动进行时间的情形, 并提出算法 SCOUT-A (Scheduling seCurity resOurces in pUblic evenTs with no relocating delAy) 来求解此情形下安保方的最优策略. 然后, 本文又研究了更具一般性的资源转移时间不可忽略的情况. 针对这种情形, 本文从离散时间假设着手, 提出算法 SCOUT-D (Scheduling seCurity resOurces in pUblic evenTs with Discrete defender strategy space) 来求解在离散时间假设下的安保方最优策略, 又在此基础上设计出算法 SCOUT-C (Scheduling seCurity resOurces in pUblic evenTs against Continuous strategy space), 以求解连续策略空间, 资源转移时间不可忽略的情形下的安保方最优策略. 最后, 本文通过实验验证了算法的有效性[1)].

2 问题描述与博弈模型

本文假设在活动中共有 n 个可能被攻击的目标, 例如马拉松比赛中的 n 个赛段, 表示为 $\mathcal{T} = \{1, \ldots, n\}$. 图 1 通过波士顿马拉松的赛道展示了一个将比赛场地分为 4 个可能被攻击的目标的样例. 假设安保方共有 m ($m \leqslant n$) 个相同的安全资源, 例如, 每个资源可能是一个巡逻队. 假设赛事在时刻 0 开始, 在时刻 $t_{\mathrm{e}} > 0$ 结束. 每个目标 $i \in \mathcal{T}$ 均对应一个重要性函数 $v_i(t)$ ($t \in [0, t_{\mathrm{e}}]$), 其在时刻 t 的函数值即为目标 i 在时刻 t 的重要程度. 图 2 基于图 1 的假设, 展示了 4 个目标所对应的重要性函数的一个样例. 在比赛刚开始时, 人群聚集在赛道的起始路段, 因此目标 1 的重要性函数值最高. 随着比赛的进行, 人群逐渐沿赛道转移, 因此目标 1 的重要性函数值逐渐下降. 相反地, 目标 4, 也即靠近终点的赛道, 重要性函数的值随时间而增加. 由于重要性函数实际上取决于赛事的地形、参赛者、观众数等基本信息, 而这类信息通常都是公开的, 因此我们假设重要性函数对于安保方与攻击方均为已知. 为了计算的便捷, 我们假设 $v_i(t)$ 为分段线性连续函数. 此类函数被大量用来近似更为复杂的函数形式 [9], 而近似的精确性能够通过设置线段数量来控制.

安保方在时刻 0 将安全资源部署在一些目标上, 随着活动的进行, 它在某些时刻将资源转移到其他的目标去. 安保方的一个纯策略 S 即包含时刻 0 的资源部署情况以及在时间段 $[0, t_{\mathrm{e}}]$ 内所做的所

1) 本文作者对重大赛事安保问题的初步研究见文献 [14]. 与之相比, 本文进一步扩展了模型, 对涉及到的定理等进行了更详尽的证明, 并增加了基于真实数据的仿真实验.

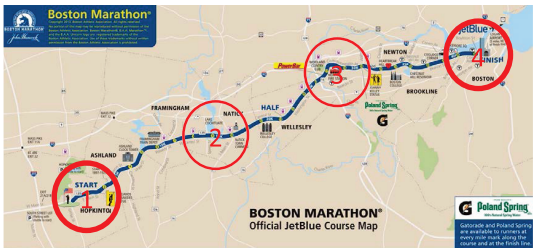

图 1 (网络版彩图) 保护及攻击目标样例

Figure 1 (Color online) Example targets

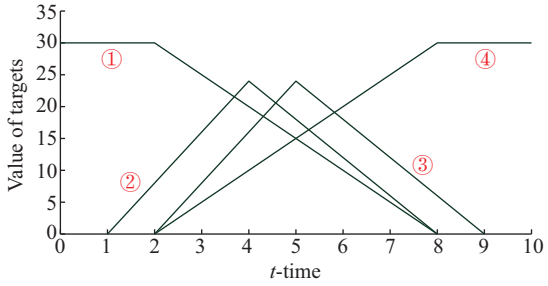

图 2 (网络版彩图) 重要性函数样例

Figure 2 (Color online) Example value functions

有转移. 我们用向量 $C=\langle C_k\rangle$ 来表示所有转移, 其中 $C_k=\langle c_{ij}^k : i,j\in\mathcal{T}\rangle$ 表示第 k 次转移, c_{ij}^k 代表第 k 次转移中从目标 i 到目标 j 转移的资源的个数. 我们用一个向量 $D=\langle d_{ij} : i,j\in\mathcal{T}\rangle$ 来表示转移资源所需的时间, 其中 d_{ij} 即为从目标 i 到目标 j 转移资源所需时间. 另外, 我们定义向量 $\tau=\langle\tau_k\rangle$, 令 τ_k 表示第 k 次转移发生的时间. 令 $Q^0=\langle q_i^0\rangle$ 表示在赛事开始时安保资源的部署情况, 其中 q_i^0 表示赛事开始时目标 i 被分配的资源个数. 相似地, 令 $Q^t(S)=\langle q_i^t(S) : i\in\mathcal{T}\rangle$ 表示任意时刻 t 安全资源的部署情况. 那么, 一个安保策略即可被完整的表示为 $S=(Q^0,C,\tau)$. 而 $q_i^t(S)$ 可以通过如下公式得到.

$$q_i^t(S)=q_i^0+\sum_{C_k\in C,\tau_k\leqslant t-d_{ji},j\in\mathcal{T}}c_{ji}^k-\sum_{C_k\in C,\tau_k\leqslant t,j\in\mathcal{T}}c_{ij}^k. \tag{1}$$

攻击方的一个纯策略是在时间段 $[0,t_{\mathrm{e}}]$ 内选择一个时间点 t, 并选择一个目标 i, 在这个时间点对这个目标进行攻击, 可表示为一个元组 (i,t).

值得注意的是, 某个目标在某个时刻可能被多个安全资源保护, 而不同数量的安全资源对于目标的保护程度是不同的. 令 $p(r)$ 表示当有 r 个资源保护某个目标时, 若攻击方选择攻击此目标, 其攻击能够成功的概率. 常规来讲, $p(r)$ 应满足以下两个条件.

(1) $p(r)\in[0,1]$, 且当 $r=0$ 时, $p(r)=1$ (即无保护时攻击总能成功), 当 $r=\infty$ 时 $p(r)=0$ (即当有无穷多资源保护某目标时攻击不可能成功).

(2) $\frac{\partial p(r)}{\partial r}\leqslant 0$ 并且 $\frac{\partial^2 p(r)}{\partial r^2}\geqslant 0$, 即经济学上的边际递减效应 —— 随着资源数目的增加, 再增加一个资源带来的边际收益减少.

为满足这两个条件, 令 $p(r)=\frac{1}{\mathrm{e}^{\lambda r}}(\lambda>0)$, 其中 λ 是用来衡量增加一个资源带来的收益的参数. 那么, 假设在时刻 t, 目标 i 被 r 个资源保护. 那么若攻击方在此刻发动攻击, 其收益即为 $W_i^r(t)=p(r)v_i(t)$, 类似地, 给定安保方的一个策略 S, 假设攻击方的策略是 (i,t), 那么攻击方可获得的收益即可表示为 $U^a(i,t,S)=W_i^{q_i^t(S)}(t)$. 可想而知, 安保方的目的是和攻击方相反的, 因此我们考虑零和博弈, 即给定双方的策略, 安保方的收益为 $U^d(i,t,s)=-U^a(i,t,s)$.

由于安保方与攻击方之间进行的是一次性零和博弈, 在均衡条件下, 安保方的策略应满足使其最坏收益最好, 也即, 即使攻击方能够对安保策略做出最优回应, 安保方的策略也能为其带来最优收益. 假设给定某个安保策略 S, 攻击方能够选择的最优策略是 $f(S)=\{f_{\mathrm{tg}}(S):S\to i, f_{\mathrm{tm}}(S):S\to t\}$, 其中 $f_{\mathrm{tg}}(S)$ 是其选择攻击的目标, 而 $f_{\mathrm{tm}}(S)$ 是其选择攻击的时间. 那么一组均衡策略 $(S,f(S))$ 应满足如下条件:

$$U^a(f_{\mathrm{tg}}(S),f_{\mathrm{tm}}(S),S)\geqslant U^a(i,t,S),\quad \forall i\in\mathcal{T},t\in[0,t_{\mathrm{e}}], \tag{2}$$

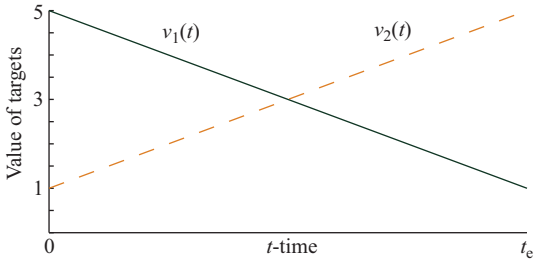

图 3 (网络版彩图) 重要性函数

Figure 3 (Color online) Value functions

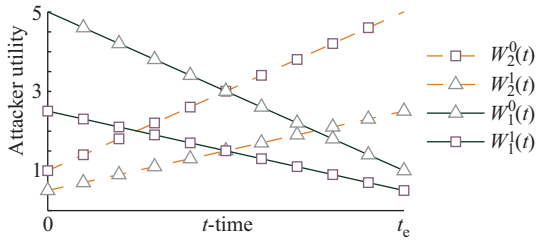

图 4 (网络版彩图) 可能的攻击方收益

Figure 4 (Color online) Possible attacker utilities

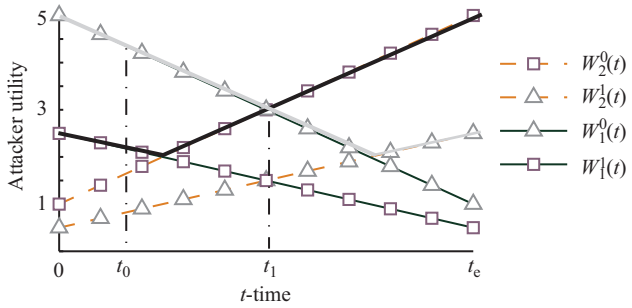

图 5 (网络版彩图) 最优配置方案

Figure 5 (Color online) Minimax assignment

$$U^d(f_{\rm tg}(S), f_{\rm tm}(S), S) \geqslant U^d(f_{\rm tg}(S'), f_{\rm tm}(S'), S'), \quad \forall S' \in \mathcal{S}. \tag{3}$$

不等式 (2) 限定给定安保方的策略 S, 攻击方做出最优的选择. 不等式 (3) 则限定即使攻击方做出最优选择, 安保方的均衡策略仍应使其获得最优的收益. 接下来, 我们介绍算法来求解这样的安保方策略.

3 转移时间可忽略的情形

首先从一个简单的情形开始研究这个问题, 即, 安保资源的转移时间可以忽略不计. 这种情形对应的现实场景即为各个目标距离较近, 因而资源的转移时间远小于赛事的进行时间. 当转移时间为 0 时, 一个直观的求解安保方最优策略的方法即求出在 $[0, t_{\rm e}]$ 内的每个时刻安保方的最优资源配置方案. 我们将一个配置方案定义为 $A = \langle a_i : \sum_{i\in\mathcal{T}} a_i = m\rangle$, 其中 a_i 表示目标 i 所配置的资源数量. 令 $\mathcal{A}$ 表示所有资源配置方案的集合. 对于任何一个安保策略 S, 在零转移时间的假设下, 任何一个时刻所有资源均在使用中, 故 $Q^t(S) \in \mathcal{A}$. 以下示例可直观展示此种情形.

例1 两个目标的重要性函数 $v_i(t)$ 如图 3 所示. 安保方有一个资源. 故而 $\mathcal{A} = \{A = \langle 1, 0\rangle, A' = \langle 0, 1\rangle\}$. 如果安保方在时刻 t 使用配置方案 $\langle 1, 0\rangle$, 那么攻击方在任一时刻攻击两个目标所能够获得的收益, 即 $W_1^1(t)$ 和 $W_2^0(t)$, 可被图 4 中以方块标示的线条标示. 类似地, 以三角标示的线条显示当安保方采用配置方案 $\langle 0, 1\rangle$ 时攻击方可能获得的收益. 我们的目标即计算何时使用方案 $\langle 1, 0\rangle$, 何时使用方案 $\langle 0, 1\rangle$, 以使得攻击方在 $[0, t_{\rm e}]$ 时间段内可能获得的最大攻击收益最小化.

定义配置方案 $A = \langle a_i\rangle$ 是时刻 t 的最优配置, 仅当对于任何 $A' = \langle a'_i\rangle \in \mathcal{A}$, 有 $\max_{i\in\mathcal{T}} W_i^{a_i}(t)$

$\leqslant \max_{i\in\mathcal{T}} W_i^{a_i'}(t)$. 考虑例 1 中的情形, $\max_{i\in\mathcal{T}} W_i^{a_i}(t)(A=\langle 1,0\rangle)$ 由图 5 中黑色粗线条表示, 灰色粗线条表示 $\max_{i\in\mathcal{T}} W_i^{a_i'}(t)(A'=\langle 0,1\rangle)$. 不难发现, 在时刻 t_0, $\max_{i\in\mathcal{T}} W_i^{a_i}(t_0) \leqslant \max_{i\in\mathcal{T}} W_i^{a_i'}(t_0)$, 所以 A 是时刻 t_0 的最优配置方案. 下述命题描述了一组最优策略.

命题1 如果一个安保策略 S 满足 $\forall t\in[0,t_{\mathrm{e}}]$, $Q^t(S)$ 是在时刻 t 的最优配置, 那么 S 是攻击方的最优策略.

证明 由于 $Q^t(S)$ 是时刻 t 的最优配置, 那么对于 $S'\in\mathcal{S}$ 和 $\forall t\in[0,t_{\mathrm{e}}]$, 有

$$\max_{i\in\mathcal{T}} W_i^{q_i^t(S)}(t) \leqslant \max_{i\in\mathcal{T}} W_i^{q_i^t(S')}(t). \tag{4}$$

于是,

$$\begin{aligned} U^a(f_{\mathrm{tg}}(S),f_{\mathrm{tm}}(S),S) &= \max_{i\in\mathcal{T}} W_i^{q_i^t(S)}(t) \quad (t=f_{\mathrm{tm}}(S)) \\ &\leqslant \max_{i\in\mathcal{T}} W_i^{q_i^t(S')}(t) \quad (t=f_{\mathrm{tm}}(S)) \\ &\leqslant U^a(f_{\mathrm{tg}}(S'),f_{\mathrm{tm}}(S'),S'). \end{aligned}$$

也即 S 使得攻击方的最大收益最小.

命题 1 说明可以通过求出每个时刻的最优配置来求解安保方的最优策略. 本文关注最优配置的一个子集 —— 完全最优配置方案, 这种配置方案总是存在且利于计算. 其定义如下.

定义1 配置方案 A 是一种在时刻 t 的完全最优配置 (简称 PA), 如果其满足 $\max_{j\in\mathcal{T}} W_j^{a_j}(t) \leqslant W_i^{a_i-1}(t)=\frac{v_i(t)}{e^{\lambda(a_i-1)}}, \forall i\in\mathcal{T}, a_i>0$.

直观上讲, 如果一种资源的配置方法与 PA 相比, 只有一个资源的分配情况不同, 那么 PA 比这种资源配置的方法更优.

命题2 一个时刻 t 的完全最优配置 A 是该时刻的最优配置.

证明 给定 $A=\langle a_i\rangle$, 任何一种配置 $A'=\langle a_i'\rangle$ 可被表示为 $a_i'=a_i+k_i$ ($\forall i\in\mathcal{T}$), 其中 k_i 是一个整数. 由于资源总数固定, 有 $\sum_{i\in\mathcal{T}} k_i=0$. 如果 $A'\neq A$, 那么必然存在至少一个目标 $j\in\mathcal{T}$, 满足 $k_j<0$. 于是有 $W_j^{a_j'}(t)=\frac{v_j(t)}{e^{\lambda(a_j+k_j)}} \geqslant \frac{v_j(t)}{e^{\lambda(a_j-1)}}=W_j^{a_j-1}(t)$. 如果 A 是时刻 t 的完全最优配置, 那么 $W_j^{a_j-1}(t)\geqslant \max_{i\in\mathcal{T}} W_i^{a_i}(t)$, 于是 $W_j^{a_j'}(t)\geqslant \max_{i\in\mathcal{T}} W_i^{a_i}(t)$. 由于 $\max_{i\in\mathcal{T}} W_i^{a_i'}(t)\geqslant W_j^{a_j'}(t)$, 则有 $\max_{i\in\mathcal{T}} W_i^{a_i'}(t)\geqslant \max_{i\in\mathcal{T}} W_i^{a_i}(t)$.

那么, 我们就可以通过计算每个时刻 $t\in[0,t_{\mathrm{e}}]$ 的 PA 来求解安保方的最优策略. 然而, 由于时间的连续性, 列举每个时刻并求解其 PA 在计算上并不可行. 但可以证明, 虽然时间是连续的, 每个时刻对应的 PA 并不会连续变化, 那么我们就只需计算有限个 PA. 假设配置 A 是时刻 t_0 的 PA. 对于任何一对目标 i,j, 如果它们满足 $\exists t\in[t_0,t_{\mathrm{e}}]$, 就有 $\frac{\partial W_i^{a_i-1}(t)}{\partial t} < \frac{\partial W_j^{a_j}(t)}{\partial t}$. 令 $I_{ij}(A,t_0)$ 表示线 $W_i^{a_i-1}(t)$ 和线 $W_j^{a_j}(t)$ 的第一个交点. 以下命题说明了 PA 有效的时间段.

命题3 如果配置 A 在时刻 t_0 是 PA, 那么 A 在任何一个时刻 $t\in[t_0,\min(\boldsymbol{I}(A,t_0))]$ 均为 PA.

证明 用反证法来证明该命题. 令 $t_1=\min(\boldsymbol{I}(A,t_0))$. 如果存在一个时刻 $t_2\in(t_0,t_1)$ 使得在此时刻 A 不是一个完全最优配置, 那么即存在一对目标 i,j 使得 $W_i^{a_i-1}(t_2)<W_j^{a_j}(t_2)$. 由于 A 是 t_0 时刻的完全最优配置, 有 $W_i^{a_i-1}(t_0)\geqslant W_j^{a_j}(t_0)$. 那么, 线 $W_i^{a_i-1}(t)$ 和线 $W_j^{a_j}(t)$ 必然在某个时刻 $t\in[t_0,t_2]$ 处相交. 那么 $t_1>t_2\geqslant\min(\boldsymbol{I}(A,t_0))$, 这与我们在 t_1 时刻的假设不符.

由于 A 是 t_0 时刻的完全最优配置, 那么有 $\max_{j\in\mathcal{T}} W_j^{a_j}(t_0)<W_i^{a_i-1}(t_0)$ ($\forall i\in\mathcal{T}, a_i>0$). 如果不存在目标对 i,j 在任一时刻 $t\in[t_0,t_{\mathrm{e}}]$ 满足 $\frac{\partial W_i^{a_i-1}(t)}{\partial t} < \frac{\partial W_j^{a_j}(t)}{\partial t}$, 那么 $\max_{j\in\mathcal{T}} W_j^{a_j}(t)\leqslant W_i^{a_i-1}(t)$ ($\forall i\in$

$\mathcal{T}, t \in [t_0, t_e]$). 那么 A 直到 t_e 都是完全最优配置. 类似地, 如果不存在目标对 i, j 使得线 $W_i^{a_i-1}(t)$ 和线 $W_j^{a_j}(t)$ 在时间段 $[t_0, t_e]$ 内相交, A 直到 t_e 都是完全最优配置.

命题 3 说明, 欲求解安保方的最优策略, 我们只需考虑那些 PA 发生变化的时间点, 并求解这些时间点上的新 PA. 再次考虑例 1 中的情形. 图 5 显示, 在时刻 0, 由于 $W_1^0(0) > \max\{W_1^1(0), W_2^0(0)\}$, $A = \langle 1, 0\rangle$ 是最优配置. t_1 是线 $W_1^0(t)$ 和线 $W_2^0(0)$ 的交点, 也是集合 $\boldsymbol{I}(A, 0)$ 中唯一的元素. 所以 A 在 $[0, t_1]$ 内均为 PA. 可以从图 5 中直接观测到, 在时刻 t_1 之后, A 即不再是完全最优配置. 接下来, 探索如何在 PA 发生变化的时刻求解新的 PA.

命题4 如果配置 A 在时刻 t 是 PA, 并且存在目标对 i, j 使得 $W_i^{a_i-1}(t) = W_j^{a_j}(t)$, 那么如果一个配置 A' 满足 $a'_i = a_i - 1, a'_j = a_j + 1, a'_k = a_k$ $(\forall k \in \mathcal{T}, k \neq i, j)$, A' 也是时刻 t 的 PA.

证明 由于只有分配给目标 i 的资源数目减少, 分配给目标 j 的资源数目增加, 而其他目标所享有的资源数目保持不变. 为证明 A' 是完全最优配置, 只需证明 $W_j^{a'_j-1}(t) \geqslant W_k^{a'_k}(t)$ 和 $W_k^{a'_k-1}(t) \geqslant W_i^{a'_i}(t) (\forall k \in \mathcal{T}, a'_k > 0)$. 首先, 由于 $W_j^{a'_j-1}(t) = W_j^{a_j}(t) = W_i^{a_i-1}(t)$ 以及 $W_i^{a_i-1}(t) \geqslant W_k^{a'_k}(t)$, 有 $W_j^{a'_j-1}(t) \geqslant W_k^{a'_k}(t)$. 类似地, 由于 $W_k^{a'_k-1}(t) \geqslant W_j^{a_j(t)}$ 和 $W_j^{a_j(t)} = W_i^{a_i-1}(t) = W_i^{a'_i}(t)$, 有 $W_k^{a'_k-1}(t) \geqslant W_i^{a'_i}(t)$.

算法 1 SCOUT-A

for all $i \in \mathcal{T}$ **do**
  $V_i \leftarrow v_i(0), a_i \leftarrow 0$
end for
left $\leftarrow m$
while left > 0 **do**
  $i \leftarrow \text{argmax}_{i\in\mathcal{T}} V_i$, $a_i \leftarrow a_i + 1$, left $\leftarrow$ left $- 1$, $V_i \leftarrow W_i^{a_i}(0)$
end while
$t_m \leftarrow 0$, $k \leftarrow 0$
while $t_m < t_e$ **do**
  $\boldsymbol{I} \leftarrow \emptyset$
  for all $\forall i, j \in \mathcal{T}$ **do**
    if $\exists t$ such that $\frac{\partial W_i^{a_i-1}(t)}{\partial t} < \frac{\partial W_j^{a_j}(t)}{\partial t}$ **then**
      $I_{ij} \leftarrow I_{ij}(A, t_m)$
      if $I_{ij} < t_e$ **then**
        $\boldsymbol{I} \leftarrow \boldsymbol{I} \cup I_{ij}$
      end if
    end if
  end for
  if $\boldsymbol{I} = \emptyset$ **then**
    break
  end if
  $I_{ij} \leftarrow \min(\boldsymbol{I})$
  if $W_i^{a_i-1}(I_{ij} - \Delta t) \geqslant W_j^{a_j}(I_{ij} - \Delta t)$ **then**
    $\tau_k \leftarrow I_{ij}, t_m \leftarrow I_{ij}, a_i \leftarrow a_i - 1, a_j \leftarrow a_j + 1, c_{ij}^k \leftarrow 1, k \leftarrow k + 1$
  end if
end while

图 5 显示, 在例 1 中, $A = \langle 1, 0\rangle$ 在时刻 t_1 是 PA. 由于 $W_1^{a_1-1}(t_1) = W_1^0(t_1) = W_2^{a_2}(t_1) = W_2^0(t_1)$, $A' = \langle 0, 1\rangle$ 在时刻 t_1 同样是 PA. 对于 A', 在 t_1 之后, 没有目标对 i, j 满足 $\frac{\partial W_i^{a_i-1}(t)}{\partial t} < \frac{\partial W_j^{a_j}(t)}{\partial t}$. 所以

直到 t_{e}, A' 都是 PA.

SCOUT-A 算法 (算法 1) 利用了上文中命题所讨论的性质, 其主要思想如下: (1) 在当前时间点计算 PA. (2) 计算此 PA 失效的时间点. (3) 将该时间点设置为当前时间点. 重复这 3 步直到对于所有目标对 i, j, 线 $W_i^{a_i-1}(t)$ 和线 $W_j^{a_j}(t)$ 在当前时间点到 t_{e} 内都不相交.

现在逐行解释算法 1 所示的 SCOUT-A 算法. 第 1∼7 行计算时刻 0 的 PA. 首先, 所有的目标都没有被保护 (第 1 行). 然后, 资源被逐个分配给当前具有最高攻击方收益的目标 (第 4∼7 行). 第 8 行中的 t_m 用来记录当前 PA 失效的时间点. k 用来记录转移的次数. 第 9∼22 行计算下一次转移的时间点. 如果在时刻 t_{e} 之前, 任何 $W_i^{a_i-1}(t)$ 与 $W_j^{a_j}(t)$ 均无交点, 当前 PA 直到 t_{e} 都是 PA, 算法结束 (第 19, 20 行), 否则 I_{ij} 被用于记录其交点. 第 23∼25 行计算新的 PA. 在第 23 行中, Δt 是一个比任何相邻的两个 $W_i^{a_i}(t)$ 和 $W_j^{a_j}(t)$ $(\forall i, j \in \mathcal{T}, \forall a_i, a_j \in \{0, \ldots, m\})$ 的交点都要小的值. 第 23 行检测是否该配置在 I_{ij} 之后仍为 PA. 这是为了避免算法在某个时刻循环交换两种配置. 由于 SCOUT-A 所求出的安保策略在任何时间点均为最优配置, 因此该策略也即安保方的最优策略.

4 转移时间不可忽略的情形

本节讨论更一般的转移时间不能忽略的情形. 在此情形下, 在某个时间点 t 可能有大于等于 1 个资源正在转移途中, 因此每个时间点在使用中的资源数目不固定, SCOUT-A 就不能用来求解此情形. 接下来, 从较简单的离散策略空间入手, 进而推进到最具一般性的转移时间不可忽略, 连续策略空间的情况.

4.1 离散策略空间

首先假设安保方只能在既定的, 离散的时间点进行资源的转移, 设时间点的集合为 $\Phi = \{t_k\}$. 这样, 一个资源到达任何一个目标的时间只可能在集合 $\phi = \{t_\delta : t_\delta = t_k + d_{ij}, \forall t_k \in \Phi, \forall i, j \in \mathcal{T}\}$ 中. 于是, 可以用一个有序的向量 $\Psi = \{t_\eta : t_\eta \in \Phi \quad \text{or} \quad t_\eta \in \phi\}$ 来表示所有可能的有资源被转移或有资源到达某个新目标的时间点. 令 η 表示 Ψ 中元素的索引值, 如, $t_\eta \in \Psi$. 令 $H = |\Psi|$. 对于 $\eta \in \{1, \ldots, H\}$, 定义一个向量 $\sigma_\eta = \langle \sigma_{ij}^\eta : \sigma_{ij}^\eta \in \{1, \ldots, H\} \rangle$, 其中 σ_{ij}^η 表示如果一个资源从目标 i 转移到目标 j, 而且预计到达 j 的时刻为 t_η, 那么转移应该在时刻 $t_{\sigma_{ij}^\eta}$ 开始. 我们引入两个新变量来表示某个时刻资源的分配情况: $a_i^{t_\eta}$, 代表在时刻 t_η, 尚没有资源被从目标 i 转移走时, 目标 i 所拥有的资源数目; $b_i^{t_\eta}$, 代表在时刻 t_η, 所有预计在此刻开始从目标 i 转移的资源均已移出后, 目标 i 所剩下的资源数目. 那么, 安保方的最优策略就可以通过如下的混合线性规划求得. 我们将此混合线性规划命名为 SCOUT-D.

$$\min U \tag{5}$$

$$\text{s.t.} \quad \sum\nolimits_{i \in \mathcal{T}} a_i^0 = m, \tag{6}$$

$$a_i^{t_{k+1}} = b_i^{t_k} + \sum\nolimits_{j \in \mathcal{T}} c_{ji}^{\sigma_{ji}^{k+1}}, \quad \forall k \in \{1, \ldots, H-1\}, \tag{7}$$

$$b_i^{t_k} = a_i^{t_k} - \sum\nolimits_{j \in \mathcal{T}} c_{ij}^k, \quad \forall k \in \{1, \ldots, H\}, \tag{8}$$

$$c_{ij}^k \in \{0, 1, \ldots\}, \quad \forall i, j \in \mathcal{T}, \forall k \in \{1, \ldots, H\}, \tag{9}$$

$$\sum c_{ij}^{t_\eta} = 0, \quad \forall t_\eta \in \phi \text{ 和 } t_\eta \notin \Phi, \tag{10}$$

$$\sum c_{ij}^{\sigma_{ij}^{\eta}}=0,\quad \forall t_\eta\in\Phi \text{ 和 } t_\eta\notin\phi, \tag{11}$$

$$b_i^{t_k}\geqslant 0\quad \forall i\in\mathcal{T},\ \forall k\in\{1,\ldots,H\}, \tag{12}$$

$$U\geqslant \max_{t\in[t_k,t_{k+1}]} W_i^{b_i^{t_k}}(t),\quad \forall i\in\mathcal{T},\ \forall k\in\{1,\ldots,H-1\}. \tag{13}$$

方程 (6) 保证了初始资源配置的有效性, 即资源总数为 m 个. 在方程 (7) 中, $\sum_{j\in\mathcal{T}} c_{ji}^{\sigma_{ji}^{k+1}}$ 代表在时刻 t_{k+1} 到达目标 i 的资源数目. 类似地, 在方程 (8) 中, $\sum_{j\in\mathcal{T}} c_{ij}^{k}$ 代表在时刻 t_k 由目标 i 转移到其他目标的资源数. 方程 (9) 限定了转移资源的数量应该为整数. 方程 (7)∼(9) 共同限制了资源转移的可行性, 类似于网络流问题中对于流可行性的限制. 方程 (10) 处理那些属于 ϕ 但不属于 Φ 的时间点, 在这些时间点安保方不能开始转移资源. 类似地, 方程 (11) 处理那些属于 Φ 但不属于 ϕ 的时间点, 在这些时间点没有资源会到达新的目标. 方程 (13) 用于限制攻击方做出最优反应, 此处攻击方可以在任何时间进行攻击. 由于 $v_i(t)$ 是连续函数, $\max_{t\in[t_k,t_{k+1}]} W_i^{b_i^{t_k}}(t)$ 总是存在, 并可以被事先计算出作为此混合线性规划的输入. 目标函数和方程 (13) 共同保障了即使攻击方做出最优的回应, 安保方仍能取得最优的收益.

4.2 连续策略空间

安保方事实上可以在时间段 $[0,t_{\mathrm{e}}]$ 内的任一时刻开始进行资源转移. 值得注意的是, 并不是任何时刻的转移都会带来好的收益, 并且, 如果把一些转移提前或延迟到另一些时刻进行, 安保方的收益可能并不会变化. 也就是说, 有可能可以通过合理的设置时间集 Φ, 使得在上一节中提出的混合线性规划求出的解同样是拥有无限策略空间的安保方的最优解. 本节首先证明, 对于任何一个安保方拥有无限策略空间的博弈, 必然存在一个与之等价的博弈, 其中安保方只能在某些既定的时间点进行资源转移. 然后, 提出一个算法来求解这些时间点, 那么所求得的时间集即可作为混合线性规划的输入, 以求出原博弈 (即安保方拥有无限空间的博弈) 的最优解.

首先, 将重要性函数分割成一系列的单调区间. 对于每个目标 i, 令 $\xi_1^i,\xi_2^i,\ldots,\xi_{R_i}^i$ 按序的表示所有函数 $v_i(t)$ 的单调性发生变化的时刻. 令 $\xi_0^i=0$ 并且 $\xi_{R_i+1}^i=t_{\mathrm{e}}$. 定义一个集合 $\Xi^i=\{\xi_\rho^i:\rho\in\{0,\ldots R_i+1\}\}$, 其中 ρ 是用来表示某个特定时间点的索引, $\xi_\rho^i\in\Xi^i$. 那么, 在任意两个在 Ξ^i 中相邻的时间点所代表的时间段内, $v_i(t)$ 是单调的. 定义 $\Xi=\{\Xi^i:\forall i\in\mathcal{T}\}$.

如果说, 在一个安保方具有无限策略空间的博弈中, 任何一个安保方的最优策略 S, 都可以被转化成另一个同样是最优的策略, 其中转移只开始在某些特定的时间点, 那么就可以先求出这些特定的时间点, 并将其作为混合线性规划的输入. 显然, 在此情况下, 混合线性规划返回的最优解也即原始博弈的最优解. 接下来, 讨论如何将原博弈的最优策略 S 进行转化. 首先, 可以证明, 将一些转移进行合并并不会影响安保方的收益.

定理1 如果在策略 S 中, 一个资源在时刻 t_1 被从目标 i 转移到目标 j, 随后在时刻 t_2 被从目标 j 转移到目标 l, 如果 $t_1+d_{ij}\in[\xi_\rho^j,\xi_{\rho+1}^j)$ 并且 $t_2\in[t_1+d_{ij},\xi_{\rho+1}^j)$, 那么如果合并这两次转移, 即, 直接在时刻 $t_3\in[t_1,t_2+d_{jl}-d_{il}]$ 将资源从目标 i 转移到目标 l, 安保方的收益不会降低.

证明 假设在初始策略 S 中, 在 t_1 时刻的转移发生前, 分布在目标 i,j,l 上的资源数量分别是 a_i,a_j,a_l. 那么在策略 S, 这 3 个目标拥有的资源数量随时间变化的情况即如表 1 所示.

值得注意的是, 应有 $d_{il}\leqslant d_{ij}+d_{jl}$, 否则安保方可将 l 作为中转站来从目标 i 向目标 l 转移资源. 于是, 总是存在时间点 $t_3\in[t_1,t_2+d_{jl}-d_{il}]$. 现在构造一个安保策略 S^1. 在 S^1 中, 除资源直接在 t_3

表 1　策略 S 中目标 i,j,l 的资源数量

Table 1　Resources assigned to i,j,l in S

Time period	$q_i^t(S)$	$q_j^t(S)$	$q_l^t(S)$
Before t_1	a_i	a_j	a_l
$[t_1, t_1+d_{ij})$	a_i-1	a_j	a_l
$[t_1+d_{ij}, t_2)$	a_i-1	a_j+1	a_l
$[t_2, t_2+d_{jl})$	a_i-1	a_j	a_l
After t_2+d_{jl}	a_i-1	a_j	a_l+1

表 2　策略 S^1 中目标 i,j,l 的资源数量

Table 2　Resources assigned to i,j,l in S^1

Time period	$q_i^t(S)$	$q_j^t(S)$	$q_l^t(S)$
Before t_3	a_i	a_j	a_l
$[t_3, t_3+d_{il})$	a_i-1	a_j	a_l
After t_3+d_{il}	a_i-1	a_j	a_l+1

时刻从目标 i 转到目标 l 之外, 其他均与 S 相同. 那么, 目标 i,j,l 在 S^1 中的资源数量随时间变化情况即如表 2 所示.

由表 1 和 2 可知, 只有在时间段 $[t_1+d_{ij}, t_2)$ 内, S^1 中分配给目标 j 的资源数量少于 S 中分配给 j 的资源数量. 其他的任何时刻, S^1 中任何目标所拥有的资源数均不少于其在 S 中拥有的资源数. 也就是说, 仅当 $\max_{t\in[t_1+d_{ij},t_2)} W_j^{q_t^j(S^1)}(t) > U^a(f_{\rm tg}(S), f_{\rm tm}(S), S)$ 时, S^1 带来的安保方收益才会小于 S. 然而, 由于 $v_j(t)$ 在时间段 $[\xi_\rho^j, \xi_{\rho+1}^j]$ 内单调, 那么 $\max_{t\in[t_1+d_{ij},t_2)} W_j^{a_j}(t) \leqslant \max\{W_j^{a_j}(t_1+d_{ij}), W_j^{a_j}(t_2)\} \leqslant U^a(f_{\rm tg}(S), f_{\rm tm}(S), S)$. 那么合并转移以后的策略 S^1 所带来的安保方收益不会少于 S.

基于定理 1, 可以将一个原博弈中的任一最优策略 S 转化为一个新的最优策略 S^1. 接下来, 证明可以通过替换部分 S^1 中的转移得到一个新的最优策略, 其中转移只发生在特定的时间点. 假设在 S^1 中, 一个资源在时刻 $t_1\in[\xi_\rho^i, \xi_{\rho+1}^i)$ 被从目标 i 转移出, 并在时刻 $t_2\in[\xi_{\rho'}^j, \xi_{\rho'+1}^j)$ 到达目标 j. 假设在 t_1 时刻之前不久, 目标 i 所拥有的资源数目为 a_i. 假设在 t_2 时刻之前不久, 目标 j 所拥有的资源数量为 a_j. 那么 $\mathrm{Tr}=(i,j,a_i,a_j,\rho,\rho')$ 即可表示一次转移. 为了探索如何替换 Tr, 首先定义一个和 Tr 相关的特定时间点 $\theta(\mathrm{Tr})$.

$$\theta(\mathrm{Tr}) = \arg \min_{t\in[\xi_\rho^i, \min\{\xi_{\rho+1}^i, \xi_{\rho'+1}^j - d_{ij}\}]} \Big(\max_{t'\in[t, t+d_{ij}]} \{W_i^{a_i-1}(t'), W_j^{a_j}(t')\} \Big). \tag{14}$$

$\theta(\mathrm{Tr})$ 的含义是, 已知目标 i 所拥有资源数量为 a_i, 目标 j 所拥有资源数量为 a_j, 如果一个资源在时间段 $[\xi_\rho^i, \xi_{\rho+1}^i]$ 被从目标 i 转出, 并在时间段 $[\xi_{\rho'}^j, \xi_{\rho'+1}^j]$ 到达目标 j, 那么如果将此转移改为在时刻 $\theta(\mathrm{Tr})$ 开始, 如果攻击方选择在此资源转移途中攻击目标 i 或 j, 那么攻击方所获得的最大收益将被最小化. 基于 $\theta(\mathrm{Tr})$ 的定义, 我们有如下关于 S^1 的转化的定理.

定理2　如果将 S^1 中所有转移 Tr 的开始时间都移至 $\theta(\mathrm{Tr})$, 安保方的收益不变.

证明　假设在策略 S^1 中, 转移 Tr 的开始时间为 t_1. 我们构造一个策略 S^2, 其中转移 Tr 开始的时间均被移至 $\theta(\mathrm{Tr})$. 接下来, 在 $t_1<\theta(\mathrm{Tr})$ 的假设下证明定理 2. 如果 $t_1\geqslant\theta(\mathrm{Tr})$, 证明过程相似. 当 $t_1<\theta(\mathrm{Tr})$, 仅当 $t\in(t_1+d_{ij}, \theta(\mathrm{Tr})+d_{ij})$, S^2 中分配给目标 j 的资源数比 S^1 少一个. 对于其他任何目标 k, 在任何时间点 t, 均有 $q_k^t(S^2)\geqslant q_k^t(S^1)$. 那么 $\max_{t\in(t_1+d_{ij},\theta(\mathrm{Tr})+d_{ij})} W_j^{q_t^j(S^2)}(t) >$

$U^a(f_{\mathrm{tg}}(S^1), f_{\mathrm{tm}}(S^1), S^1)$ 时, S^1 才会比 S^2 带来更高的安保方收益.

给定 $\theta(\mathrm{Tr})$ 的定义, 可知 $\theta(\mathrm{Tr}) + d_{ij} \in [\xi^j_{\rho'}, \xi^j_{\rho'+1}]$. S^1 的性质保证了在时间段 $[t_1+d_{ij}, \xi^j_{\rho'+1}]$ 内, 没有资源被从目标 j 转出. 令 ω 代表时间段 $[t_1 + d_{ij}, \theta(\mathrm{Tr}) + d_{ij}]$, 那么当 $t \in \omega$, 有 $q^t_j(S^2) \geqslant a_j$. 于是, $\max_{t\in\omega} W_j^{q^j_t(S^2)}(t) \leqslant \max_{t\in\omega} W_j^{a_j}(t)$. 为证明 $\max_{t\in\omega} W_j^{a_j}(t) \leqslant U^a(f_{\mathrm{tg}}(S^1), f_{\mathrm{tm}}(S^1), S^1)$, 将以下两种情况分开讨论.

(1) $t_1 + d_{ij} \geqslant \theta(\mathrm{Tr})$. 在此情况下, 有

$$\max_{t\in\omega} W_j^{a_j}(t) \leqslant \max_{t\in[\theta(\mathrm{Tr}),\theta(\mathrm{Tr})+d_{ij}]} W_j^{a_j}(t) \leqslant \max_{t\in[\theta(\mathrm{Tr}),\theta(\mathrm{Tr})+d_{ij}]} \{W_i^{a_i-1}(t), W_j^{a_j}(t)\}.$$

基于 $\theta(\mathrm{Tr})$ 的定义, 可知

$$\max_{t\in[\theta(\mathrm{Tr}),\theta(\mathrm{Tr})+d_{ij}]} \{W_i^{a_i-1}(t), W_j^{a_j}(t)\} \leqslant \max_{t\in[t_1,t_1+d_{ij}]} \{W_i^{a_i-1}(t), W_j^{a_j}(t)\} \leqslant U^a(f_{\mathrm{tg}}(S^1), f_{\mathrm{tm}}(S^1), S^1).$$

于是 $\max_{t\in\omega} W_j^{q^j_t(S^2)}(t) \leqslant U^a(f_{\mathrm{tg}}(S^1), f_{\mathrm{tm}}(S^1), S^1)$, 也即策略 S^2 带来的安保方收益不少于策略 S^1.

(2) $t_1 + d_{ij} < \theta(\mathrm{Tr})$. 如果 $\max_{t\in[t_1+d_{ij},\theta(\mathrm{Tr})]} W_j^{a_j}(t) \leqslant \max_{t\in[\theta(\mathrm{Tr}),\theta(\mathrm{Tr})+d_{ij}]} W_j^{a_j}(t)$, 那么不难推知

$$\max_{t\in\omega} W_j^{a_j}(t) \leqslant \max_{t\in[\theta(\mathrm{Tr}),\theta(\mathrm{Tr})+d_{ij}]} W_j^{a_j}(t),$$

那么就可通过情况 1 的方法证明 $\max_{t\in\omega} W_j^{a_j}(t) \leqslant U^a(f_{\mathrm{tg}}(S^1), f_{\mathrm{tm}}(S^1), S^1)$. 否则, 由于函数 $v_j(t)$ 在时间段 $[\xi^j_{\rho'}, \xi^j_{\rho'+1}]$ 内单调, 当 $t \in [\xi^j_{\rho'}, \xi^j_{\rho'+1}]$, 有 $\frac{\mathrm{d}W_j^{a_j}(t)}{\mathrm{d}t} \leqslant 0$. 那么,

$$\max_{t\in\omega} W_j^{a_j}(t) \leqslant W_j^{a_j}(t_1 + d_{ij}) \leqslant U^a(f_{\mathrm{tg}}(S^1), f_{\mathrm{tm}}(S^1), S^1).$$

接下来, 描述对于任意一个安保方有连续无限策略空间的博弈, 求解其最优解的算法. 首先, 定义一个集合

$$\Theta = \{\theta(\mathrm{Tr}) : \mathrm{Tr} = (i, j, a_i, a_j, \rho, \rho'), \forall i, j \in \mathcal{T}, \forall a_i, a_j \in \{0, \dots m\}, \forall \rho \in \{0, \dots, R_i\}, \forall \rho' \in \{0, \dots, R_j\}\}.$$

定理 2 显示, 可将任一安保策略 S^1 转化为另一个策略 S^2 而不影响安保方的最优收益, 在 S^2 中, 所有的转移仅在时间点 $t \in \Theta$ 开始. 也就是说, 只要原始博弈, 即安保方有连续无限策略空间的博弈, 存在一个最优策略 S, 那么它还必然存在一个最优策略, 其中转移仅在时间点 $t \in \Theta$ 开始. 我们提出算法 SCOUT-C (算法 2) 来求解这一最优策略. SCOUT-C 首先计算出在一个最优策略中所有转移可能开始的时间点和所有转移可能结束的时间点 (第 3 行). 这些时间点被记录在一个集合 Ψ 中. 接下来, Ψ 被用作混合线性规划 SCOUT-D 的输入, 而如前所述, SCOUT-D 所返回的最优策略, 同样是原始博弈, 即安保方拥有无限连续策略空间博弈的最优解.

算法 2 SCOUT-C

1: $\Psi \leftarrow \emptyset$
2: **for all** $\rho \in \{0, \dots, R_i\}, \rho' \in \{0, \dots, R_j\}$ **do**
3: $\quad \Psi \leftarrow \Psi \cup \{\theta(\mathrm{Tr})\} \cup \{\theta(\mathrm{Tr}) + d_{ij}\}$, where $\mathrm{Tr} = (i, j, a_i, a_j, \rho, \rho')$
4: **end for**
5: run SCOUT-D, using Ψ as the time points set

5 实验与分析

分别在完全随机数据和基于真实场馆分布的模拟数据上测试了文章提出的算法. 首先介绍随机数据下的实验结果. 在随机数据中, 生成重要性函数时, 对于每个目标, 我们首先在时间段 $[0, t_{\mathrm{e}}]$ 随机选

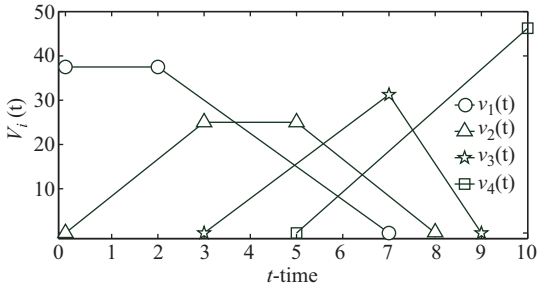

图 6　阶段线性重要性函数

Figure 6　Piecewise linear value functions

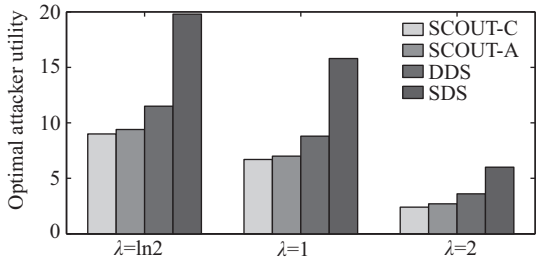

图 7　改变 λ 值

Figure 7　Changing λ

择一段, 要求其中重要性函数的值不为零. 然后随机产生该目标的重要性函数对应的线性阶段数, 再生成每个阶段对应的线性函数. 假设一个目标的重要性函数值在 $[0, 100]$ 以内. 图 6 展示了一个随机生成的 4 个目标的重要性函数示例, 其中 t_{e} 被设为 10, x 轴表示时间, y 轴表示重要性函数的值.

首先检测在不同的 λ 值和不同的转移时间下, 算法对于安保方收益的影响. 为便于展示, 用攻击方的收益来衡量安保方收益. 由于安保方收益与攻击方收益相反, 越低的攻击方收益对应越高的安保方收益. 为了更好地衡量本文算法的有效性, 同时展示了另外两个算法, SDS (static defender strategy) 和 DDS (dynamic defender strategy), 所带来的攻击方收益. 在 SDS 中, 安保方使用静态策略, 即, 每个目标被分配的资源数目与其重要性函数的最大值成正比, 在赛事进行的整个过程中, 资源不发生转移. 类似的静态策略被广泛应用于之前关于安全博弈的研究工作 [15~17]. 在 DDS 中, 假设时间被时间分割为离散的点 $\Phi = \{\frac{n \cdot t_{\mathrm{e}}}{4} : n \in \{0, 1, \ldots, 4\}\}$, 并用 SCOUT-D 求解在此假设下的安保方最优策略.

图 7 展示了在不同的 λ 值下, SCOUT-A, SCOUT-C, DDS 和 SDS 所带来的攻击方收益. 横轴对应 3 种 λ 值, 纵轴对应攻击方收益. 随着 λ 值的增大, 4 种算法带来的攻击方收益都减少, 也即安保方收益都增加, 这是由于越大的 λ 值说明安保资源的保护效力越大, 安保方更容易获得好的收益. 值得注意的是, 不论 λ 取何值, SCOUT-C 和 SCOUT-A 所对应的攻击方收益都明显小于 DDS, 而 SDS 带来的攻击方收益都是 4 种算法中最高的. 这说明 SCOUT-C 和 SCOUT-A 所产生的安保策略要优于 DDS, 而动态的安保策略又要优于静态的安保策略.

图 8 展示了在不同的转移时间级别下 4 种算法所带来的攻击方收益. 横轴对应 3 种不同等级的转移时间: 对于 Level 1, 任何两个目标之间的转移时间, 即 d_{ij}, 被认为是在 $[0, 0.1]$ 内均匀分布; 对于 Level 2, d_{ij} 被认为在 $[0, 1]$ 内均匀分布; 对 Level 3, d_{ij} 则被认为在 $[0, 5]$ 内均匀分布. 纵轴表示 4 种算法所带来的攻击方收益. 无论转移时间在何种级别, SCOUT-C 都带来最低的攻击方收益, SDS 都带来最差的安保策略. 图 9 展示了在不同的转移时间范围内, SCOUT-A 与 SCOUT-C 带来的攻击方收益的差异. 如图 8, 横轴对应 3 种转移时间等级, 而纵轴表示攻击方收益的差值. 当转移时间很小时, 两种算法的差别也很小. 当转移时间增加, 两种算法的差距也变大.

我们还检测了算法的运行时间随博弈规模增加而变化的情况. 图 10 展示了两种算法的运行时间. 其中 x 轴表示 4 组目标数/资源数的组合情况, 如 8/10 表示该博弈中共有 8 个目标需要保护, 共有 10 个安全资源可供分配. y 轴为运行时间. SCOUT-A 表现出了更好的可扩展性. 图 11 展示了在更大规模的博弈上 SCOUT-A 的运行时间. 横轴表示目标数目, 3 条曲线分别表示安保资源数为 10, 20, 30. 即使目标数目达到 100 而资源数目达到 30, SCOUT-A 仍能在 30 分钟内求解出安保方的最优策略.

为了检测算法的扩展潜能, 我们考虑了形式更一般的重要性函数, 即, 重要性函数可为分段二次

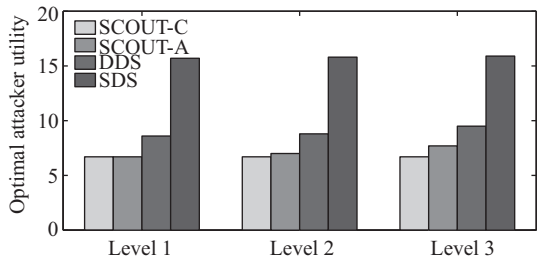

图 8 改变转移时间

Figure 8 Changing range of d_{ij}

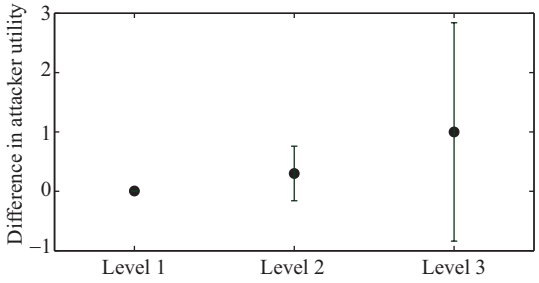

图 9 SCOUT-C 与 SCOUT-A 的差别

Figure 9 SCOUT-C v.s. SCOUT-A

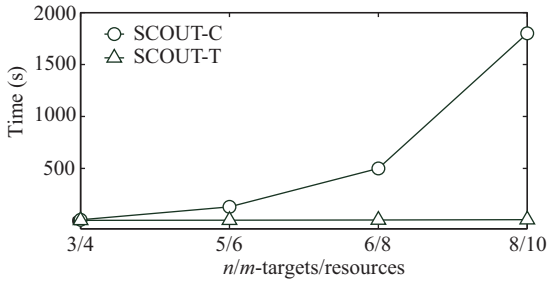

图 10 运行时间

Figure 10 Runtime

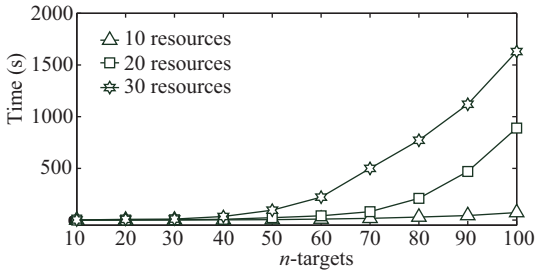

图 11 SCOUT-A 在大规模博弈上的运行时间

Figure 11 SCOUT-A on large games

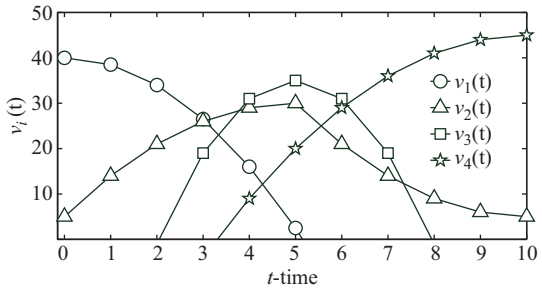

图 12 更普适的重要性函数

Figure 12 General value functions

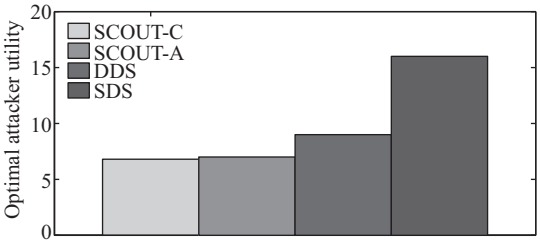

图 13 攻击方收益

Figure 13 Attacker's utility

函数. 图 12 展示了随机生成的 4 个目标的阶段二次函数形式的重要性函数. 图 13 展示了在此类函数形式下, SCOUT-A, SCOUT-C, SDS 和 DDS 分别对应的攻击者收益情况. 结果显示, 即使在更一般的重要性函数形式下, SCOUT-C 和 SCOUT-A 依然带来比 SDS 和 DDS 更优的安保策略.

最后, 在真实场馆的地形分布上测试了算法的效果. 图 14 展示了即将到来的 2022 年北京冬奥会中, 北京城区的 5 个主要场馆的分布情况. 表 3 展示了在这些场馆之间转移所需要的时间. 由于当前比赛日程并未公布, 我们仍使用随机生出的数据来模拟各场馆的重要性变化情况. 假设所有场馆的赛事进行时间均为 10 小时, 以分钟为时间单位, 即 $t_{\mathrm{e}} = 600$. 同时, 以随机试验中所采用的方法为各个场馆生成阶段线性的重要性函数. 假设有 10 个巡逻队可在这 5 个场馆间机动转移. 图 15 展示了 4 种算

图 14 (网络版彩图) 北京冬奥会主要场馆 (图片来自申奥委官网)

Figure 14 (Color online) Studios for 2022 Olympic Winter Games

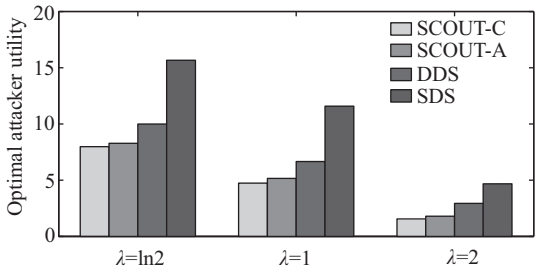

图 15 攻击方收益

Figure 15 Attacker utility

表 3 场馆间转移时间 (min, 来自谷歌地图的估算)

Table 3 Transfer time between studios (min, from google map)

	鸟巢	奥林匹克公园	首体	工体	五棵松
鸟巢	–	8	14	17	29
奥林匹克公园	8	–	20	24	33
首体	14	20	–	24	17
工体	17	24	24	–	36
五棵松	29	33	17	36	–

法所带来的收益情况. 类似于图 7, 横轴表示 λ 的数值, 纵轴表示攻击方的最优收益. SCOUT-C 的表现仍然明显优于 SDS 和 DDS. 事实上, 由于场馆间的转移时间远小于赛事的总进行时间, SCOUT-C 所求得的攻击方收益仅仅略差于不考虑转移时间的 SCOUT-A 的效果.

6 结论

为重大赛事设计高效的安保策略是具有重大现实意义的议题. 然而, 由于在重大赛事中, 可能被攻击的目标的重要程度是随时间变化的, 而可行的安保策略以及安保方可能面对的攻击策略都数量巨大, 因此设计最优的安保策略在技术上极具挑战性. 本文用博弈模型来描述重大赛事的安保问题, 在模型中既考虑了目标重要性的动态变化, 也考虑了攻击方与安保方连续无限的策略空间. 基于此模型, 本文考虑了转移时间可忽略不计, 转移时间不可忽略不计的两种情况. 针对第一种情况, 本文提出算法 SCOUT-A 求解安保方的最优收益. 针对第二种情况, 本文从离散时间点着手, 提出了在此假设下

求解最优安保策略的算法 SCOUT-D, 进而考虑了更为复杂的连续时间的情形, 提出了基于 SCOUT-D 求解最优安保策略的算法 SCOUT-C. 实验结果表明, SCOUT-A 和 SCOUT-C 所产生的安保策略均优于此前研究中不考虑目标重要性动态变化时所生成的安保策略.

参考文献

1 Tambe M. Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned. Cambridge: Cambridge University Press, 2011

2 An B, Kempe D, Kiekintveld C, et al. Security games with limited surveillance. In: Proceedings of the 26th AAAI Conference on Artificial Intelligence, Toronto, 2012. 1241–1248

3 Shieh E, An B, Yang R, et al. PROTECT: an application of computational game theory for the security of the ports of the United States. In: Proceedings of the 26th AAAI Conference on Artificial Intelligence, Toronto, 2012. 2173–2179

4 Agmon N, Urieli D, Stone P. Multiagent patrol generalized to complex environmental conditions. In: Proceedings of the 25th AAAI Conference on Artificial Intelligence, San Francisco, 2011. 1090–1095

5 Basilico N, Gatti N, Amigoni F. Leaderfollower strategies for robotic patrolling in environments with arbitrary topologies. In: Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), Budapest, 2009. 57–64

6 Letchford J, Vorobeychik Y. Optimal interdiction of attack plans. In: Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), Saint Paul, 2013. 199–206

7 Basilico N, Gatti N, Amigoni F. Patrolling security games: definition and algorithms for solving large instances with single patroller and single intruder. Artif Intell, 2012, 184–185: 78–123

8 An B, Ordonez F, Tambe M, et al. A deployed quantal response-based patrol planning system for the US Coast Guard. Interfaces, 2013, 43: 400–420

9 Yang R, Ordonez F, Tambe M. Computing optimal strategy against quantal response in security games. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), Valencia, 2012. 847–854

10 Yin Z, Jiang A X, Johnson M P, et al. TRUSTS: scheduling randomized patrols for fare inspection in transit systems. In: Proceedings of the 24th Innovative Applications of Artificial Intelligence (IAAI), Toronto, 2012. 2348–2355

11 An B, Brown M, Vorobeychik Y, et al. Security games with surveillance cost and optimal timing of attack execution. In: Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), Saint Paul, 2013. 223–230

12 Fang F, Jiang A X, Tambe M. Optimal patrol strategy for protecting moving targets with multiple mobile resources. In: Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), Saint Paul, 2013. 957–964

13 Xu H, Fang F, Jiang A X, et al. Solving zero-sum security games in discretized spatio-temporal domains. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence, Québec City, 2014. 1500–1506

14 Yin Y, An B, Jain M. Game theoretic resource allocation for protecting large public events. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence, Québec City, 2014. 826–834

15 Kiekintveld C, Islam T, Kreinovich V. Security games with interval uncertainty. In: Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), Saint Paul, 2013. 231–238

16 Korzhyk D, Conitzer V, Parr R. Solving Stackelberg games with uncertain observability. In: Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), Taipei, 2011. 1013–1020

17 Varakantham P, Lau H C, Yuan Z. Scalable randomized patrolling for securing rapid transit networks. In: Proceedings of the 25th Innovative Applications of Artificial Intelligence Conference (IAAI), Bellevue, 2013. 1563–1568

Designing game-theoretic security strategies for large public events

Yue YIN[1,2*], Bo AN[1,3] & Zhongzhi SHI[1]

1. *The Key Lab of Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing* 100190*, China*;
2. *University of Chinese Academy of Sciences, Beijing* 100049*, China*;
3. *School of Computer Science and Engineering, Nanyang Technological University, Singapore* 639798*, Singapore*
* Corresponding author. E-mail: melody1235813@163.com

Abstract High-profile, large-scale public events may be attractive targets for terrorist attacks. The security challenge for such events is exacerbated by their dynamic nature: the impact of an attack on different 'targets', such as studio entrances, changes over time. In addition, the defender can relocate security resources among potential attack targets at any time, while the attacker may act at any time during the event. This study focuses on developing efficient patrolling algorithms for such dynamic domains, with continuous strategy spaces for both the defender and attacker. We propose SCOUT-A, which makes assumptions regarding relocation costs, exploits payoff representation, and computes optimal solutions efficiently. We furthermore propose SCOUT-C, to compute the exact optimal defender strategy for general cases despite the continuous strategy spaces. The experimental results demonstrate that our algorithms significantly outperform existing strategies.

Keywords game theory, security, algorithms, strategy design, multi-agent

Yue YIN was born in 1989. She received her B.S. degree in information security from the Beijing Information Science and Technology University. She is currently a Ph.D. candidate at the Institute of Computing Technology, Chinese Academy of Sciences. Her research interests lie mainly in multi-agent systems, game theory, and social choice.

Bo AN was born in 1980. He received a Ph.D. degree in Computer Science from the University of Massachusetts, Amherst. He is a Nanyang assistant professor at the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests include artificial intelligence, multi-agent systems, game theory, automated negotiation, resource allocation, and optimization.

Zhongzhi SHI was born in 1941. He graduated from the Graduate University of Chinese Academy of Sciences in 1968. He is currently a professor at the Institute of Computing Technology, Chinese Academy of Sciences, where he leads the Intelligence Science Laboratory. His major research and teaching interests are in the areas of intelligence science, distributed intelligence, machine learning, neural computing, and data mining.